

04 Caso de estudio / Diario de Colón

Paul Spence, King's College London 2026

Introducción

Resumen

Esta es la actividad principal del taller de esta semana. Primero analizaremos un texto de humanidades y luego lo marcaremos en XML en varias etapas con Oxygen.

Métodos de aprendizaje/Apuntes de laboratorio

Este ejercicio combinará debate con toda la clase y trabajo en grupo. Les sugerimos que tomen notas de laboratorio a medida que trabajan en las diferentes partes de los ejercicios, anotando lo aprendido, las dificultades que encuentren y cualquier pregunta (técnica o teórica) que tengan para más adelante.

El texto

El texto con el que trabajaremos es una traducción al inglés de una versión del diario de Cristóbal Colón (que era un diario) del sitio web del proyecto Early Modern Spain, editado por el profesor Barry Ife: <https://ems.kcl.ac.uk/content/etext/e019-copyright.html>

Si les interesa, consulten Wikipedia para obtener más información:
https://es.wikipedia.org/wiki/Diario_de_la_primera_navegaci%C3%B3n

Etapas 1 – Análisis

Al analizar un texto (lo que a veces se denomina "análisis de documentos"), lo primero que hay que hacer es considerar aspectos generales y de alto nivel del texto, como:

- **Metadatos:** se trata de información general sobre el documento, como el título, el autor y el proceso de edición (desde su creación hasta la actualidad, incluyendo las acciones realizadas en el documento). Los metadatos son "datos sobre datos", el tipo de información que se puede utilizar para facilitar el descubrimiento y que puede aparecer en la portada.
- **Estructura general.** ¿Cuáles son los principales elementos estructurales del texto? Si se analiza un libro de cocina, por ejemplo, se podría analizar de la siguiente manera:

Libro de cocina

Metadatos

Título

Autor
Fecha de publicación
Dificultad
etc.
Contenido
Capítulos
Título del capítulo
Receta
Nombre de la receta
Foto (opcional)
Ingredientes
Instrucciones de la receta

- Tenga en cuenta que estos elementos tienen una jerarquía, que se muestra aquí mediante sangría. Al dibujar la estructura del diario de Colón, puede que quiera mostrarla mediante líneas u otro mecanismo visual.
- Puede ser útil **dividirlo en pasos**, por ejemplo:
 - concentrarse primero en las *estructuras de alto nivel*
 - y luego abordar las *estructuras a nivel de fragmentos*, como párrafos y listas.

Etapa 1 – Marcado

Puntos generales

- Como antes, **tomen notas** mientras realizan esta actividad. Guarden notas sobre aspectos prácticos en una sección y anoten reflexiones más generales (por ejemplo, sobre los posibles beneficios o limitaciones del marcado) en otra. ¿Qué funcionó bien? ¿Qué posibilidades ven para esta forma de anotar un texto?
- Trabajen en grupos, pero es mejor que todos marquen, ya que se les pedirá que dividan el trabajo más adelante.
- **No sientan la necesidad de marcar todo el texto; prioricen la experimentación** sobre la exhaustividad en este ejercicio.
- **Guarden su progreso regularmente y guarden copias**, o creen diferentes versiones del XML si desean ver las diferentes etapas más adelante.
- Una vez que se sientan cómodos con el análisis de la estructura general, estarán listos para comenzar a marcar el documento en XML.

Creen la plantilla de marcado

- No lo hagan en orden cronológico estricto, sino **en orden lógico**, por ejemplo. Usando el ejemplo del recetario:
- Primero, elija el elemento de nivel superior: <cookbook>
- Luego, elija los elementos que aparecen debajo: <metadatos> y <contenido>
- etc.

Agregue el contenido

- Una vez que esté satisfecho con la estructura general, comience a **copiar el contenido**.

- No se preocupe si a veces la estructura no funciona correctamente; es normal tener que ajustar el esquema de marcado al compararlo con el contenido. Sin embargo, **tome nota de cualquier cambio que realice** en esta etapa; esto es muy útil para la fase de discusión.

Cosas para recordar

- Guarde el archivo en cuanto comience.
- Compruebe que el archivo esté correctamente formateado.

Etapa 2 - Análisis

- Los elementos que identificamos en la primera etapa eran (en su mayoría) jerárquicos. Por ejemplo, un libro de cocina tiene metadatos, contenido, capítulos y una serie de recetas. Más abajo en la jerarquía, una receta tiene un nombre, una foto (opcional), una serie de ingredientes y un método.
- **¿Qué ocurre con los elementos no jerárquicos/no estructurales?** Siguiendo el mismo ejemplo, ¿sería útil buscar métodos de cocina, por ejemplo, o visualizar términos culinarios? Para un libro de cocina, esto es un poco forzado (!), pero para algo como el diario de Colón podría ser útil buscar **términos**, y estos términos podrían no ser fáciles de encontrar mediante una búsqueda no estructurada.
- **Haz una lista de los términos que podrías querer buscar en las entradas del diario.** Piensa en lo que necesitarías capturar para cada uno de estos términos. Si elegiste información personal como término a capturar, por ejemplo, piensa en lo que realmente estás capturando.
- **¿Estás "marcando" el nombre, la persona o ambos?** Para que la distinción sea más clara, si capturas "elena gonzalez" como nombre, estás marcando a todas las personas con ese nombre, pero un programa no podrá distinguir entre personas con el mismo nombre.

¿Y **cómo se marcaría esto** exactamente? ¿Se marcaría toda la cadena de texto como un nombre o se codificaría el nombre y el apellido por separado?

Etapa 2 - Marcado

Acuerden un esquema de marcado

- Acuerden **los nombres que usarán para cada elemento**
- Intente usar atributos en dos o tres elementos

Marque el texto

- Una vez que estén listos, continúen **marcando el texto**. Les sugerimos dividir el trabajo para que puedan avanzar más y ayudarse mutuamente a resolver cualquier duda.
- El extracto que les proporcionamos incluye:
 - Una introducción seguida de cuatro entradas cortas
 - Entrada del 11 de octubre
 - Entrada del 19 de octubre

- Entrada del 21 de diciembre
 - Dos entradas cortas seguidas de material adicional
- Les sugerimos dividir el marcado siguiendo estas líneas (o similares).
- Una vez que todos hayan terminado, **comparen su marcado**:
 - Comprueben que todos usaron exactamente los mismos nombres de elemento: <Receta> no es lo mismo que <receta>
 - ¿Marcaron elementos diferentes?
 - ¿Usaron el mismo marcado de forma diferente? Si tuvieras tiempo, ¿cómo ampliarías tu marcado?

Conclusiones y reflexiones

- **Reflexiona**
- Al empezar a marcar, estás tan ocupad* aprendiendo a usar la herramienta que te resulta difícil reflexionar.
- Reflexiona sobre el trabajo realizado y piensa cómo lo mejorarías si tuvieras más tiempo.
- **¿Cómo ampliarías tu modelo de marcado para poder editar una colección más amplia de materiales históricos?**
- **¿Cómo imaginas usar el marcado** para crear nuevas visualizaciones o nuevas funcionalidades?

Otros puntos a tener en cuenta

- Normalmente no empezarías a marcar un texto tan pronto; es mejor dedicar tiempo a considerar la historia, el estado y el contexto de un documento o corpus.
- También deberías considerar a tu público. ¿Qué público académico podría estar interesado en este documento? ¿Qué disciplinas representan y cómo podría cada disciplina plantear diferentes preguntas sobre el texto? ¿Cómo podría interesar el texto a un público más amplio? ¿Qué, específicamente, podría desear cada uno de estos públicos hacer con el texto? ¿Qué tipo de visualizaciones o funciones interpretativas podrían utilizar para estudiar el texto y cómo podrían ayudar las herramientas digitales?
- En un proyecto real, es probable que se utilice un estándar preestablecido como DITA o TEI en lugar de crear uno propio.

Referencia - Edición de XML con Oxygen

Acerca de este documento

- Este documento está pensado como una referencia rápida para otras áreas de edición de XML en Oxygen. Es bastante básico; para una explicación más completa, consulte los excelentes tutoriales disponibles en otros lugares:
- Tutoriales generales de XML como <https://www.w3schools.com/xml/> (en inglés)
- Tutoriales de humanidades digitales como <https://teach.dariah.eu/course/view.php?id=23> (en inglés)
- Tutoriales específicos sobre el uso de Oxygen en <https://www.youtube.com/@oxygenxml> (en inglés). La mayoría de los tutoriales van mucho más allá del alcance de este módulo,

pero algunos de los siguientes pueden ser interesantes para ver cómo se utiliza el editor para la creación de contenido en entornos comerciales, por ejemplo:

- *Basics of Oxygen XML for Technical Writers- Part 1*
- *The XML Authoring videos*

Acerca del editor Oxygen

- Le solicitamos que descargue el editor Oxygen completo, que está destinado a desarrolladores. Existen otras versiones, por ejemplo: Solo para autoría, pero nos gustaría que pudieras ver algunas de las opciones de desarrollo.
- El editor Oxygen tiene tres vistas (Text | Grid | Author - Texto | Cuadrícula | Autor). Te recomendamos usar la versión Texto, que muestra todas las etiquetas.
- Observa que la sintaxis está coloreada para facilitar la lectura. Los elementos están en azul, la instrucción de procesamiento XML en la línea 1 está en morado y el contenido del texto está en negro. Esto no forma parte de XML; Oxygen añade los colores para facilitar el trabajo.
- Observa que aparece un pequeño recuadro verde en la esquina superior derecha al iniciar un nuevo documento. Esto indica que la sintaxis del archivo XML es correcta. En la jerga XML, está bien formado.
- Si el documento no está bien formado, por ejemplo, Si borras una etiqueta de cierre por error, el cuadro se vuelve rojo.
- Oxygen proporciona un manual si deseas obtener información más detallada sobre cómo usar el editor: <https://www.oxygenxml.com/doc/versions/22.1/ug-editor/topics/getting-started-first-XML-document.html> (hay que crear cuenta)

Cómo insertar el texto en el documento XML

- En general, existen dos opciones:
- Copiar y pegar todo el texto y luego agregar el XML.
- Crear primero la estructura XML principal y luego copiar y pegar el texto fragmento a fragmento.
- Recomendamos la segunda opción, que es menos propensa a errores y más fácil de corregir si surgen problemas.

Elementos XML

- Los elementos en XML se componen de:
- etiquetas de inicio, p. ej., <book>
- luego contenido
- luego una etiqueta de cierre, p. ej., </book>. Ten en cuenta la barra diagonal para cerrar el elemento, p. ej., <book>.

```
<book>Contenido del libro aquí</book>
```

- Cada documento XML debe tener un elemento que contenga todo el resto del contenido.
- El elemento XML puede contener texto sin formato u otro contenido, p. ej.:

```
<title>100 años</title>
```

- o

```
<book><title>100 años</title></book>
```

- Los elementos XML no pueden superponerse, p. ej.:

```
<title><emphasis>Texto</emphasis></title>
```

- no

```
<title><emphasis>Texto</title></emphasis>
```

Atributos

- Se puede usar solo elementos en un archivo XML, pero normalmente se necesitan atributos.
- Los atributos pertenecen a elementos individuales y proporcionan información adicional sobre ese elemento.
- Se ven así:

```
<lugar n="Pamplona">
```

- Este es solo el elemento de apertura; el uso completo sería algo así:

```
<place n="Pamplona">la capital de Navarra</place>
```

- Como se puede ver, hay un elemento llamado <place>, que contiene:
 - un nombre de atributo: 'n'
 - un valor de atributo: 'Pamplona'
- Los aspectos clave a tener en cuenta son:
 - Hay **un espacio después del elemento**.
 - El nombre del atributo **va seguido de**
 - **un signo igual (=)**.
 - y luego el valor del atributo entre comillas.
- El atributo se añade solo a la etiqueta de apertura, no a la de cierre.
- Al igual que los elementos, **no debe haber espacios en el nombre del atributo**; por ejemplo, esto no vale <lugar nombre histórico=""> por el espacio entre 'nombre' e 'histórico'

Más información sobre los atributos

- Puede tener más de un atributo por elemento.
- Simplemente añada un espacio entre cada uno; por ejemplo,

```
<place type="city" n="Pamplona">
```

- Para más información, consulte W3schools:
https://www.w3schools.com/xml/xml_attributes.asp

Material

- Ver <https://polito66.github.io/>